

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:00:49 ; Search time 53 Seconds

(without alignments)
357.183 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 67

Sequence: 1 GFGGLGGRGKCFSEIPSR.....CRGLYLRNKKVCVPRSKG 67

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 775050

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Listing first 45 summaries

- Database : A Geneseq_29Jan04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	67	5	ABB08330 Bee venom
2	10	14.9	10	3	AAY69218 N-termina
3	9	13.4	11	3	AAY69212 Peptide f
4	8	11.9	8	3	AAY69211 Peptide f
5	8	11.9	9	3	AAY69216 Tryptic p
6	8	11.9	10	3	AAY69219 N-termina
7	7	10.4	59	3	AAG61636 Arabidops
8	6	9.0	6	3	AAG57565 Arabidops
9	6	9.0	6	3	AAY69214 Peptide f
10	6	9.0	6	3	AAY69210 Peptide u
11	6	9.0	8	2	AAW54524 High affi
12	6	9.0	8	3	AAY69217 Tryptic p
13	6	9.0	10	3	AAY69220 N-termina
14	6	9.0	10	4	AAG94716 Human com
15	6	9.0	10	4	AAG95858 Human com
16	6	9.0	10	4	AAG94765 Human com
17	6	9.0	17	4	AAU68072 Human Bre
18	6	9.0	17	4	AAU68135 Human Bre
19	6	9.0	17	4	AAU68089 Human Bre
20	6	9.0	18	3	AAU62603 Partial s
21	6	9.0	35	4	AAU60793 Androcton
22	6	9.0	35	5	AAU83468 Rat phosp
23	6	9.0	41	4	AAU88665 Human imm
24	6	9.0	42	4	AAU10256 Human pol
25	6	9.0	46	4	AAU09488 Human pol

26	6	9.0	48	4	AAU91473 Human imm
27	6	9.0	53	3	AAU56683 Human pro
28	6	9.0	53	4	AAU47907 Propionib
29	6	9.0	53	6	ABM44426 Propionib
30	6	9.0	55	4	AAU87031 Human imm
31	6	9.0	57	4	AAU00690 Human pol
32	6	9.0	59	4	AAU80752 Human hae
33	6	9.0	59	6	ABP79324 N. gonorr
34	6	9.0	63	6	ABM65670 Propionib
35	6	9.0	65	4	AAU86610 Novel hum
36	6	9.0	65	7	ADB5944 Connectiv
37	5	7.5	5	2	AAU23464 V beta 6
38	5	7.5	5	3	AAU69213 Peptide f
39	5	7.5	5	3	AAU69213
40	5	7.5	5	4	AAU03345 HBAAG-lys
41	5	7.5	5	4	AAU51147 Arylsulfo
42	5	7.5	5	4	AAU65469 Conserved
43	5	7.5	6	2	AAU23495 V beta 6
44	5	7.5	7	1	AAU70889 Plasmid p
45	5	7.5	7	3	AAU24135 Plasmodu
					ABG77734 Targettin

ALIGNMENTS

RESULT 1

ABB08330

ID ABB08330 standard; protein; 67 AA.

- XX AC ABB08330;
- XX AC
- DT 18-JUN-2002 (first entry)
- XX DT
- DE Bee venom protein Api m 6.01.
- XX DE
- KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
- XX KW
- KW Bee venom hypersensitivity; antibody; protein purification; Api m 6.01;
- XX KW
- XX immunotherapy; allergen.
- XX OS
- XX Apis sp.
- XX OS
- PN WO200188085-A2.
- XX PN
- PD 22-NOV-2001.
- XX PD
- PF 16-FEB-2001; 2001WO-IB001736.
- XX PF
- PR 18-FEB-2000; 2000US-00506978.
- XX PR
- XX (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.
- XX PA
- XX Spertini F;
- XX PI
- XX WPI; 2002-082988/11.
- XX DR

New bee venom polypeptides, useful for modulating immune responses e.g. in individual hypersensitive to the venom and for identifying individual at risk for bee venom hypersensitivity.

Example 2; Page 26; 32pp; English.

The present sequence is that of one the four isoforms of Api m 6, designated Api m 6.01. The sequence represents the central amino acid sequence shared by all four isoforms (see ABB08331, ABB08332 and ABB08333). The specification describes a substantially pure polypeptide, Api m 6, derived from bee venom and found in four isoforms. The proteins of the invention have immunosuppressant activity and may form the basis of a vaccine. Api m 6 is useful for modulating an immune response, i.e. as an allergen for immunotherapy. The protein is useful for identifying an individual at risk for bee venom hypersensitivity. The method comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific

CC for Abi m 6 are useful for purifying the protein

XX Sequence 67 AA;

Query Match 100.0%; Score 67; DB 5; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.4e-64;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRGKCPSEIFSRCDGRQFCNPVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

Db 1 GGFGLGGRGKCPSEIFSRCDGRQFCNPVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

QY 61 VPRSKCG 67

Db 61 VPRSKCG 67

RESULT 2

AAAY69218

ID AAY69218 standard; peptide; 10 AA.

XX

AC AAY69218;

DT 30-MAY-2000 (first entry)

DE N-terminal sequence of honey bee venom PX3.101 protein.

Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
metastatic cancer; Alzheimer's disease; wound healing; aging process;
antigen.

OS Apis mellifera.

XX

PN GB2341389-A.

PD 15-MAR-2000.

PF 13-SEP-1999; 95GB-00021605.

PR 14-SEP-1998; 98US-0100172P.

PA (PANP-) PAN PACIFIC PHARM INC.

PI Chi X, Lu Y;

WPI; 2000-185368/17.

Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
treating autoimmune and inflammatory disorders such as rheumatoid
arthritis.

Example 2; Page 52; 83pp; English.

The present sequence is derived from the protein PX3.101, which is a
honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
a variety of enzymes (e.g. cyclooxygenases, lipoxigenases, phospholipases
and proteases) associated with inflammatory diseases. The nucleic acids
may be used for the recombinant production of PX3.101 proteins either in
vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
culture). The nucleic acids may also be used as probes to identify
similar sequences in samples. The PX3.101 protein may be used for the
treatment of inflammatory diseases, cancers, autoimmune diseases, pain
and/or diseases associated with chemokine (especially IL-8) imbalances
such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
metastatic cancer and Alzheimer's disease in humans. It is also disclosed
that the proteins may be used to accelerate wound healing, reduce several
aging processes and protect against ultraviolet light. The proteins may

CC also be used as antigens in the production of antibodies specific for
PX3.101. The antibodies may be used as diagnostic agents to detect
PX3.101 protein in samples and to down regulate PX3.101 activity

XX Sequence 10 AA;

Query Match 14.9%; Score 10; DB 3; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRG 10

Db 1 GGFGLGGRG 10

RESULT 3

AAAY69212

ID AAY69212 standard; peptide; 11 AA.

XX

AC AAY69212;

DT 30-MAY-2000 (first entry)

DE Peptide fragment (residues 24-34) of the honey bee venom PX3.101.

Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
metastatic cancer; Alzheimer's disease; wound healing; aging process;
antigen.

OS Apis mellifera.

XX

PN GB2341389-A.

PD 15-MAR-2000.

PF 13-SEP-1999; 95GB-00021605.

PR 14-SEP-1998; 98US-0100172P.

PA (PANP-) PAN PACIFIC PHARM INC.

PI Chi X, Lu Y;

WPI; 2000-185368/17.

Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
treating autoimmune and inflammatory disorders such as rheumatoid
arthritis.

Example 3; Page 43; 83pp; English.

The present sequence is derived from the protein PX3.101, which is a
honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
a variety of enzymes (e.g. cyclooxygenases, lipoxigenases, phospholipases
and proteases) associated with inflammatory diseases. The nucleic acids
may be used for the recombinant production of PX3.101 proteins either in
vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
culture). The nucleic acids may also be used as probes to identify
similar sequences in samples. The PX3.101 protein may be used for the
treatment of inflammatory diseases, cancers, autoimmune diseases, pain
and/or diseases associated with chemokine (especially IL-8) imbalances
such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
metastatic cancer and Alzheimer's disease in humans. It is also disclosed
that the proteins may be used to accelerate wound healing, reduce several
aging processes and protect against ultraviolet light. The proteins may
also be used as antigens in the production of antibodies specific for
PX3.101. The antibodies may be used as diagnostic agents to detect

CC PX3.101 protein in samples and to down regulate PX3.101 activity
XX
SQ Sequence 11 AA;
Query Match 13.4%; Score 9; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFGLGGR 9
DB 3 GGFGLGGR 11
RESULT 4
AAI69211
ID AAY69211 standard; peptide; 8 AA.
XX
AC AAY69211;
XX
DT 30-MAY-2000 (first entry)
XX
DE Peptide fragment (residues 38-45) of the honey bee venom PX3.101.
XX
KW Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
metastatic cancer; Alzheimer's disease; wound healing; aging process;
antigen.
XX
OS Apis mellifera.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "unknown amino acid"
FT GB2341389-A.
XX
PD 15-MAR-2000.
XX
PF 13-SEP-1999; 99GB-00021605.
XX
PR 14-SEP-1998; 98US-0100172P.
XX
PA (PANP-) PAN PACIFIC PHARM INC.
XX
PI Chi X, Lu Y;
XX
WI; 2000-185368/17.
XX
Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
treating autoimmune and inflammatory disorders such as rheumatoid
arthritis.
XX
PS Example 3; Page 43; 83pp; English.
XX
The present sequence is derived from the protein PX3.101, which is a
honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
a variety of enzymes (e.g. cyclooxygenases, lipoxigenases, phospholipases
and proteases) associated with inflammatory diseases. The nucleic acids
may be used for the recombinant production of PX3.101 proteins either in
vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
culture). The nucleic acids may also be used as probes to identify
similar sequences in samples. The PX3.101 protein may be used for the
treatment of inflammatory diseases, cancers, autoimmune diseases, pain
and/or diseases associated with chemokine (especially IL-8) imbalances
such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
metastatic cancer and Alzheimer's disease in humans. It is also disclosed
that the proteins may be used to accelerate wound healing, reduce several
aging processes and protect against ultraviolet light. The proteins may
also be used as antigens in the production of antibodies specific for
PX3.101. The antibodies may be used as diagnostic agents to detect
PX3.101 protein in samples and to down regulate PX3.101 activity

CC
XX
SQ Sequence 8 AA;
Query Match 11.9%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 PSNEIFSR 20
DB 1 PSNEIFSR 8
RESULT 5
AAY69216
ID AAY69216 standard; peptide; 9 AA.
XX
AC AAY69216;
XX
DT 30-MAY-2000 (first entry)
XX
DE Tryptic peptide of honey bee venom PX3.101 protein.
XX
KW Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
metastatic cancer; Alzheimer's disease; wound healing; aging process;
antigen.
XX
OS Apis mellifera.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "unknown amino acid"
FT GB2341389-A.
XX
PD 15-MAR-2000.
XX
PF 13-SEP-1999; 99GB-00021605.
XX
PR 14-SEP-1998; 98US-0100172P.
XX
PA (PANP-) PAN PACIFIC PHARM INC.
XX
PI Chi X, Lu Y;
XX
WI; 2000-185368/17.
XX
Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
treating autoimmune and inflammatory disorders such as rheumatoid
arthritis.
XX
PS Example 2; Page 52; 83pp; English.
XX
The present sequence is derived from the protein PX3.101, which is a
honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
a variety of enzymes (e.g. cyclooxygenases, lipoxigenases, phospholipases
and proteases) associated with inflammatory diseases. The nucleic acids
may be used for the recombinant production of PX3.101 proteins either in
vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
culture). The nucleic acids may also be used as probes to identify
similar sequences in samples. The PX3.101 protein may be used for the
treatment of inflammatory diseases, cancers, autoimmune diseases, pain
and/or diseases associated with chemokine (especially IL-8) imbalances
such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
metastatic cancer and Alzheimer's disease in humans. It is also disclosed
that the proteins may be used to accelerate wound healing, reduce several
aging processes and protect against ultraviolet light. The proteins may
also be used as antigens in the production of antibodies specific for
PX3.101. The antibodies may be used as diagnostic agents to detect
PX3.101. The antibodies may be used as diagnostic agents to detect

CC PX3.101 protein in samples and to down regulate PX3.101 activity

XX Sequence 9 AA;

Query Match 11.9%; Score 8; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PSNEIFSR 20
Db 2 PSNEIFSR 9

RESULT 6
AAY69219
ID AAY69219 standard; peptide; 10 AA.

XX AC AAY69219;
XX DT 30-MAY-2000 (first entry)

XX N-terminal sequence of honey bee venom PX3.101 protein.

DE Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
XX CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
XX inflammatory disease; gene therapy; cancer; autoimmune disease; pain;
KW chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
KW systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
KW metastatic cancer; Alzheimer's disease; wound healing; aging process;
KW antigen.

XX Apis mellifera.

OS GB2341389-A.

PN 15-MAR-2000.

XX 13-SEP-1999; 99GB-00021605.

PF 14-SEP-1998; 98US-0100172P.

PR (PANP-) PAN PACIFIC PHARM INC.

XX Chi X, Lu Y;

PI WPI; 2000-185368/17.

DR Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
PT treating autoimmune and inflammatory disorders such as rheumatoid
PT arthritis.

XX Example 2; Page 52; 83pp; English.

XX The present sequence is derived from the protein PX3.101, which is a
CC honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
CC interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
CC a variety of enzymes (e.g. cyclooxygenases, lipoxigenases, phospholipases
CC and proteases) associated with inflammatory diseases. The nucleic acids
CC may be used for the recombinant production of PX3.101 proteins either in
CC vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
CC culture). The nucleic acids may also be used as probes to identify
CC similar sequences in samples. The PX3.101 protein may be used for the
CC treatment of inflammatory diseases, cancers, autoimmune diseases, pain
CC and/or diseases associated with chemokine (especially IL-8) imbalances
CC such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
CC lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
CC metastatic cancer and Alzheimer's disease in humans. It is also disclosed
CC that the proteins may be used to accelerate wound healing, reduce several
CC aging processes and protect against ultraviolet light. The proteins may
CC also be used as antigens in the production of antibodies specific for
CC PX3.101. The antibodies may be used as diagnostic agents to detect
CC PX3.101 protein in samples and to down regulate PX3.101 activity

SQ Sequence 10 AA;

Query Match 11.9%; Score 8; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGGLGG 8
Db 3 GFGGGLGG 10

RESULT 7
AAG61636
ID AAG61636 standard; protein; 59 AA.

XX AC AAG61636;
XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 79962.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0128785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130443P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 18-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.

PR 04-JUN-1999; 99US-0137502P.

PR 07-JUN-1999; 99US-0137724P.

PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141128P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160844P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 10.4%; Score 7; DB 3; Length 59;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGFGGLG 7
QY

Db	48 GGFGGLG 54	PR	18-JUN-1999;	99US-0139455P.
RESULT 8		PR	18-JUN-1999;	99US-0139456P.
AAGS7565		PR	18-JUN-1999;	99US-0139457P.
XX	AAGS7565 standard; protein; 59 AA.	PR	18-JUN-1999;	99US-0139458P.
XX	AAGS7565;	PR	18-JUN-1999;	99US-0139459P.
XX		PR	18-JUN-1999;	99US-0139460P.
XX		PR	18-JUN-1999;	99US-0139461P.
XX	18-OCT-2000 (first entry)	PR	18-JUN-1999;	99US-0139462P.
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 74196.	PR	18-JUN-1999;	99US-0139750P.
XX	Arabidopsis thaliana	PR	18-JUN-1999;	99US-0139763P.
XX	Protein identification; signal transduction pathway; metabolic pathway;	PR	21-JUN-1999;	99US-0139817P.
KW	hybridisation assay; Genetic mapping; gene expression control; Promoter;	PR	22-JUN-1999;	99US-0139899P.
KW	termination sequence.	PR	23-JUN-1999;	99US-0140353P.
XX	Arabidopsis thaliana.	PR	23-JUN-1999;	99US-0140354P.
OS	EP1033405-A2.	PR	24-JUN-1999;	99US-0140695P.
XX		PR	28-JUN-1999;	99US-0140823P.
XX	06-SEP-2000.	PR	29-JUN-1999;	99US-0140991P.
XX		PR	30-JUN-1999;	99US-0141287P.
XX	25-FEB-2000; 2000EP-00301439.	PR	01-JUL-1999;	99US-0141842P.
XX		PR	01-JUL-1999;	99US-0142154P.
XX	25-FEB-1999;	PR	02-JUL-1999;	99US-0142055P.
PR	05-MAR-1999;	PR	06-JUL-1999;	99US-0142390P.
PR	09-MAR-1999;	PR	08-JUL-1999;	99US-0142803P.
PR	23-MAR-1999;	PR	09-JUL-1999;	99US-0142920P.
PR	25-MAR-1999;	PR	12-JUL-1999;	99US-0142977P.
PR	29-MAR-1999;	PR	13-JUL-1999;	99US-0143542P.
PR	01-APR-1999;	PR	14-JUL-1999;	99US-0143624P.
PR	06-APR-1999;	PR	15-JUL-1999;	99US-0144008P.
PR	16-APR-1999;	PR	16-JUL-1999;	99US-0144085P.
PR	19-APR-1999;	PR	16-JUL-1999;	99US-0144086P.
PR	21-APR-1999;	PR	19-JUL-1999;	99US-0144325P.
PR	23-APR-1999;	PR	19-JUL-1999;	99US-0144331P.
PR	28-APR-1999;	PR	19-JUL-1999;	99US-0144332P.
PR	30-APR-1999;	PR	19-JUL-1999;	99US-0144333P.
PR	04-MAY-1999;	PR	19-JUL-1999;	99US-0144334P.
PR	05-MAY-1999;	PR	19-JUL-1999;	99US-0144335P.
PR	06-MAY-1999;	PR	20-JUL-1999;	99US-0144352P.
PR	07-MAY-1999;	PR	20-JUL-1999;	99US-0144632P.
PR	11-MAY-1999;	PR	20-JUL-1999;	99US-0144884P.
PR	14-MAY-1999;	PR	21-JUL-1999;	99US-0144814P.
PR	14-MAY-1999;	PR	21-JUL-1999;	99US-0145086P.
PR	14-MAY-1999;	PR	21-JUL-1999;	99US-0145088P.
PR	14-MAY-1999;	PR	22-JUL-1999;	99US-0145085P.
PR	14-MAY-1999;	PR	22-JUL-1999;	99US-0145087P.
PR	14-MAY-1999;	PR	22-JUL-1999;	99US-0145089P.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145145P.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145219P.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145278P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918P.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145919P.
PR	14-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	14-MAY-1999;	PR	28-JUL-1999;	99US-0145951P.
PR	14-MAY-1999;	PR	02-AUG-1999;	99US-0145388P.
PR	14-MAY-1999;	PR	02-AUG-1999;	99US-0145388P.
PR	14-MAY-1999;	PR	02-AUG-1999;	99US-0146389P.
PR	14-MAY-1999;	PR	03-AUG-1999;	99US-0147038P.
PR	14-MAY-1999;	PR	04-AUG-1999;	99US-0147204P.
PR	14-MAY-1999;	PR	04-AUG-1999;	99US-0147302P.
PR	14-MAY-1999;	PR	05-AUG-1999;	99US-0147192P.
PR	14-MAY-1999;	PR	05-AUG-1999;	99US-0147260P.
PR	14-MAY-1999;	PR	06-AUG-1999;	99US-0147260P.
PR	14-MAY-1999;	PR	06-AUG-1999;	99US-0147303P.
PR	14-MAY-1999;	PR	06-AUG-1999;	99US-0147416P.
PR	14-MAY-1999;	PR	06-AUG-1999;	99US-0147493P.
PR	14-MAY-1999;	PR	08-AUG-1999;	99US-0147935P.
PR	14-MAY-1999;	PR	08-AUG-1999;	99US-0148171P.
PR	14-MAY-1999;	PR	10-AUG-1999;	99US-0148319P.
PR	14-MAY-1999;	PR	11-AUG-1999;	99US-0148319P.
PR	14-MAY-1999;	PR	12-AUG-1999;	99US-0148341P.
PR	14-MAY-1999;	PR	13-AUG-1999;	99US-0148565P.
PR	14-MAY-1999;	PR	13-AUG-1999;	99US-0148684P.
PR	14-MAY-1999;	PR	16-AUG-1999;	99US-0149368P.

```

PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157855P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161484P.
PR 25-OCT-1999; 99US-0161485P.
PR 25-OCT-1999; 99US-0161486P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.

Query Match 10.4%; Score 7; DB 3; Length 59;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 G3FGGLG 7
Db 48 G3FGGLG 54

RESULT 9
RAY69214
ID RAY69214 standard; peptide; 6 AA.

XX 99US-0149175P.
AC 99US-0149426P.
XX 99US-0149722P.
DT 99US-0149723P.
XX 99US-0149929P.
DE 99US-0149902P.
XX 99US-0149930P.
XX 99US-0150566P.
KW 99US-0150884P.
KW 99US-0151065P.
KW 99US-0151066P.
KW 99US-0151080P.
KW 99US-0151080P.
KW 99US-0151303P.
XX 99US-0151438P.
XX 99US-0151930P.
OS 99US-0152363P.
XX 99US-0153070P.
PN 99US-0153758P.
XX 99US-0154018P.
PD 99US-0154039P.
XX 99US-0154779P.
PF 99US-0155139P.
XX 99US-0155486P.
PR 99US-0155659P.
XX 99US-0156458P.
PA 99US-0156596P.
XX 99US-0157117P.
PI 99US-0157753P.
XX 99US-0157855P.
DR 99US-0158029P.
XX 99US-0158232P.
PT 99US-0158369P.
PT 99US-0159293P.
PT 99US-0159294P.
XX 99US-0159295P.
PS 99US-0159329P.
XX 99US-0159330P.
CC 99US-0159331P.
CC 99US-0159637P.
CC 99US-0159638P.
CC 99US-0159584P.
CC 99US-0160741P.
CC 99US-0160767P.
CC 99US-0160768P.
CC 99US-0160770P.
CC 99US-0160814P.
CC 99US-0160815P.
CC 99US-0160980P.
CC 99US-0160981P.
CC 99US-0160989P.
CC 99US-0161484P.
CC 99US-0161485P.
CC 99US-0161486P.
CC 99US-0161359P.
CC 99US-0161360P.
CC 99US-0161361P.
CC 99US-0161920P.
CC 99US-0161992P.
CC 99US-0161993P.
CC 99US-0162142P.
XX 99US-0162143P.
SQ

Query Match 9.0%; Score 6; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PNVPVK 35
Db 1 PNVPVK 6

RESULT 10
RAY69210
ID RAY69210 standard; peptide; 6 AA.
XX RAY69210;
AC RAY69210;
```



```

XX OS Apis mellifera.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "unknown amino acid"
XX FT Misc-difference 2 /note= "unknown amino acid"
XX FT
XX PN GB2341389-A.
XX PD 15-MAR-2000.
XX PF 13-SEP-1999; 99GB-00021605.
XX PR 14-SEP-1998; 98US-0100172P.
XX PA (PANP-) PAN PACIFIC PHARM INC.
XX PI Chi X, Lu Y;
XX PI MPI; 2000-185368/17.
XX PT Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
XX PT treating autoimmune and inflammatory disorders such as rheumatoid
XX PT arthritis.
XX PS Example 2; Page 52; 83pp; English.
XX CC The present sequence is derived from the protein PX3.101, which is a
XX CC honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
XX CC interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
XX CC a variety of enzymes (e.g. cyclooxygenases lipoxigenases, phospholipases
XX CC and proteases) associated with inflammatory diseases. The nucleic acids
XX CC may be used for the recombinant production of PX3.101 proteins either in
XX CC vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
XX CC culture). The nucleic acids may also be used as probes to identify
XX CC similar sequences in samples. The PX3.101 protein may be used for the
XX CC treatment of inflammatory diseases, cancers, autoimmune diseases, pain
XX CC and/or diseases associated with chemokine (especially IL-8) imbalances
XX CC such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
XX CC lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
XX CC metastatic cancer and Alzheimer's disease in humans. It is also disclosed
XX CC that the proteins may be used to accelerate wound healing, reduce several
XX CC aging processes and protect against ultraviolet light. The proteins may
XX CC also be used as antigens in the production of antibodies specific for
XX CC PX3.101. The antibodies may be used as diagnostic agents to detect
XX CC PX3.101 protein in samples and to down regulate PX3.101 activity
XX SQ Sequence 8 AA;

Query Match 9.0%; Score 6; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PNVVVPK 35
DB 3 PNVVVPK 8

RESULT 13
AAY69220
ID AAY69220 standard; peptide; 10 AA.
XX AC
XX AAY69220;
XX DT 30-MAY-2000 (first entry)
XX DE
XX DE N-terminal sequence of honey bee venom PX3.101 protein.
XX KW Protein PX3.101; honey bee; venom; interleukin-8; IL-8; receptor; CXCR1;
XX KW CXCR2; cyclooxygenase; lipoxigenase; phospholipase; protease;
XX KW inflammatory disease; gene therapy; cancer; autoimmune disease; pain;

```

```

KW chemokine imbalance; rheumatoid arthritis; multiple sclerosis; psoriasis;
KW systemic lupus erythematosus; Crohn's disease; vasculitis; scleroderma;
KW metastatic cancer; Alzheimer's disease; wound healing; aging process;
KW antigen.
XX OS Apis mellifera.
XX PN GB2341389-A.
XX PD 15-MAR-2000.
XX PF 13-SEP-1999; 99GB-00021605.
XX PR 14-SEP-1998; 98US-0100172P.
XX PA (PANP-) PAN PACIFIC PHARM INC.
XX PI Chi X, Lu Y;
XX PI MPI; 2000-185368/17.
XX PT Isolated nucleic acids encoding the bee venom protein PX3.101 useful for
XX PT treating autoimmune and inflammatory disorders such as rheumatoid
XX PT arthritis.
XX PS Example 2; Page 52; 83pp; English.
XX CC The present sequence is derived from the protein PX3.101, which is a
XX CC honey bee venom isolated Apis mellifera. PX3.101 inhibits the binding of
XX CC interleukin-8 (IL-8) to its receptor (e.g. CXCR1 and CXCR2) and inhibits
XX CC a variety of enzymes (e.g. cyclooxygenases lipoxigenases, phospholipases
XX CC and proteases) associated with inflammatory diseases. The nucleic acids
XX CC may be used for the recombinant production of PX3.101 proteins either in
XX CC vivo (as part of a gene therapy protocol) or in vitro (as a fermentation
XX CC culture). The nucleic acids may also be used as probes to identify
XX CC similar sequences in samples. The PX3.101 protein may be used for the
XX CC treatment of inflammatory diseases, cancers, autoimmune diseases, pain
XX CC and/or diseases associated with chemokine (especially IL-8) imbalances
XX CC such as rheumatoid arthritis, multiple sclerosis, psoriasis, systemic
XX CC lupus erythematosus (SLE), Crohn's disease, vasculitis, scleroderma,
XX CC metastatic cancer and Alzheimer's disease in humans. It is also disclosed
XX CC that the proteins may be used to accelerate wound healing, reduce several
XX CC aging processes and protect against ultraviolet light. The proteins may
XX CC also be used as antigens in the production of antibodies specific for
XX CC PX3.101. The antibodies may be used as diagnostic agents to detect
XX CC PX3.101 protein in samples and to down regulate PX3.101 activity
XX SQ Sequence 10 AA;

Query Match 9.0%; Score 6; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGL 6
DB 5 GGFGGL 10

RESULT 14
AAG94716
ID AAG94716 standard; peptide; 10 AA.
XX AC
XX AAG94716;
XX DT 18-SEP-2001 (first entry)
XX DE
XX DE Human complementary peptide, SEQ ID NO: 910.
XX KW Human; complementary peptide; ligand; drug discovery; drug design.
XX OS Homo sapiens.
XX OS WO200142277-A2.
XX PN

```

```
XX 14-JUN-2001.
XX PD
XX PF 13-DEC-2000; 2000WO-GB004776.
XX XX
XX PR 13-DEC-1999; 99GB-00029464.
XX XX
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX XX
XX DR WPI; 2001-408419/43.
XX XX
XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX XX
XX Example 4; Page 169; 646pp; English.
XX XX
XX CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX XX
XX SQ Sequence 10 AA;
XX
XX Query Match 9.0%; Score 6; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 5 GLGGRG 10
XX Db |||||
XX 1 GLGGRG 6
XX
XX RESULT 15
XX AAG95858
XX ID AAG95858 standard; peptide; 10 AA.
XX AC
XX AC AAG95858;
XX XX
XX DT 18-SEP-2001 (first entry)
XX XX
XX DE Human complementary peptide, SEQ ID NO: 2052.
XX XX
XX KW Human; complementary peptide; ligand; drug discovery; drug design.
XX OS Homo sapiens.
XX XX
XX PN WO200142277-A2.
XX XX
XX PD 14-JUN-2001.
XX XX
XX PF 13-DEC-2000; 2000WO-GB004776.
XX XX
XX PR 13-DEC-1999; 99GB-00029464.
XX XX
XX PA (PROT-) PROTEOM LTD.
XX PI Roberts GW, Heal JR;
XX XX
XX DR WPI; 2001-408419/43.
XX XX
XX A set of peptide ligands consisting of specific complementary peptides to
PT proteins encoded by genes of the human genome, useful in an assay for
PT screening and identifying of one or more novel peptides which are drug
PT candidates or pro-drugs.
XX XX
XX Example 4; Page 336; 646pp; English.
XX XX
```

```
CC The invention relates to a set of complementary peptide ligands generated
CC from the human genome. The complementary peptides interact with their
CC relevant target proteins encoded in the human genome. They can be used as
CC reagents in drug discovery and as lead ligands to facilitate drug design
CC and development. The present sequence is a complementary peptide provided
CC in the specification
XX XX
XX SQ Sequence 10 AA;
```

```
XX Query Match 9.0%; Score 6; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 26;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
XX QY 5 GLGGRG 10
XX Db |||||
XX 1 GLGGRG 6
```

```
Search completed: April 1, 2004, 10:05:21
Job time : 56 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:04:19 ; Search time 23 Seconds
(without alignments)
150.389 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 67
Sequence: 1 GGFGLGGRKCPNSIFSR.....CRLGLYLRNKKVCVPRSKG 67

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 247770

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	14.9	10	4	US-09-394-630-20
2	9	13.4	11	4	US-09-394-630-12
3	8	11.9	8	4	US-09-394-630-11
4	8	11.9	9	4	US-09-394-630-18
5	8	11.9	10	4	US-09-394-630-21
6	7	10.4	47	2	US-08-871-355A-400
7	7	10.4	47	3	US-08-871-355A-400
8	7	10.4	47	4	US-09-201-945-400
9	6	9.0	6	4	US-09-394-630-9
10	6	9.0	6	4	US-09-394-630-14
11	6	9.0	10	4	US-09-394-630-19
12	6	9.0	10	4	US-09-394-630-22
13	5	7.5	5	3	US-09-026-633-2
14	5	7.5	5	3	US-08-842-306B-20
15	5	7.5	5	3	US-08-838-973B-18
16	5	7.5	5	4	US-09-394-630-13
17	5	7.5	5	4	US-08-771-212A-14
18	5	7.5	7	3	US-08-929-329-12
19	5	7.5	8	3	US-09-128-450-4
20	5	7.5	8	4	US-09-823-494-4
21	5	7.5	9	1	US-08-425-069-11
22	5	7.5	9	2	US-08-317-844B-11
23	5	7.5	9	3	US-08-963-168C-41
24	5	7.5	9	3	US-08-963-168C-42
25	5	7.5	9	3	US-09-353-976-3
26	5	7.5	9	3	US-09-353-976-5
27	5	7.5	9	3	US-09-353-976-6

28 5 7.5 9 3 US-09-258-754-67 Sequence 67, Appl
29 5 7.5 9 3 US-09-042-107-67 Sequence 67, Appl
30 5 7.5 9 4 US-09-722-250D-67 Sequence 67, Appl
31 5 7.5 9 4 US-09-671-937-3 Sequence 3, Appl
32 5 7.5 9 4 US-09-671-937-5 Sequence 5, Appl
33 5 7.5 9 4 US-09-671-937-6 Sequence 6, Appl
34 5 7.5 10 2 US-08-724-548-40 Sequence 40, Appl
35 5 7.5 10 2 US-08-724-548-41 Sequence 41, Appl
36 5 7.5 10 2 US-08-724-548-42 Sequence 42, Appl
37 5 7.5 10 2 US-08-724-548-43 Sequence 43, Appl
38 5 7.5 10 2 US-08-724-548-44 Sequence 44, Appl
39 5 7.5 10 3 US-07-978-674B-40 Sequence 45, Appl
40 5 7.5 10 3 US-07-978-674B-41 Sequence 41, Appl
41 5 7.5 10 3 US-07-978-674B-42 Sequence 42, Appl
42 5 7.5 10 3 US-07-978-674B-43 Sequence 43, Appl
43 5 7.5 10 3 US-07-978-674B-44 Sequence 44, Appl
44 5 7.5 10 3 US-07-978-674B-45 Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-09-394-630-20
; Sequence 20, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; CURRENT FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:N-terminal
; OTHER INFORMATION: sequence from Puri-#1
US-09-394-630-20

Query Match 14.9%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFGLGGRG 10
|||
Db 1 GGFGLGGRG 10

RESULT 2
US-09-394-630-12
; Sequence 12, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; CURRENT FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14

```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
; OTHER INFORMATION: fragment #62 from in-gel trypsin digestion
; NAME/KEY: MOD RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa = unsure amino acid
;
US-09-394-630-18

Query Match      11.9%; Score 8; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 PSNEIFSR 20
      |||||
Db       2 PSNEIFSR 9

RESULT 5
US-09-394-630-21
; Sequence 21, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; CURRENT FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:N-terminal
; OTHER INFORMATION: sequence from Puri-#2
;
US-09-394-630-21

Query Match      11.9%; Score 8; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGGLGG 8
      |||||
Db       3 GGFGGLGG 10

RESULT 6
US-08-637-759B-400
; Sequence 400, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
```

```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:residues 24-34
; OTHER INFORMATION: of SEQ ID NO:2 obtained through protein sequencing
;
US-09-394-630-12

Query Match      13.4%; Score 9; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGFGGLGGR 9
      |||||
Db       3 GGFGGLGGR 11

RESULT 3
US-09-394-630-11
; Sequence 11, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; CURRENT FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:residues 38-45
; OTHER INFORMATION: of SEQ ID NO:2 obtained through protein
; OTHER INFORMATION: sequencing; peptide fragment #75 from in-gel
; OTHER INFORMATION: trypsin digestion
;
US-09-394-630-11

Query Match      11.9%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      13 PSNEIFSR 20
      |||||
Db       1 PSNEIFSR 8

RESULT 4
US-09-394-630-18
; Sequence 18, Application US/09394630
; Patent No. 6395306
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yuefeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200US
; CURRENT APPLICATION NUMBER: US/09/394,630
; CURRENT FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: US 60/100,172
; PRIOR FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 400:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-400

Query Match 10.4%; Score 7; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SRCDCRC 25
Db 37 SRCDCRC 43

RESULT 7
US-08-871-355A-400
Sequence 400, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871.355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 400:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-400

Query Match 10.4%; Score 7; DB 3; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SRCDCRC 25
Db 37 SRCDCRC 43

RESULT 8
US-09-201-945-400
Sequence 400, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 400:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-400

Query Match 10.4%; Score 7; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 SRCDCRC 25
|||||

Db 37 SRCDGRC 43

RESULT 9

US-09-394-630-9

; Sequence 9, Application US/09394630

; Patent No. 6395306

; GENERAL INFORMATION:

; APPLICANT: Cui, Xiangmin

; APPLICANT: Lu, Yuefeng

; APPLICANT: Pan Pacific Pharmaceutical, Inc.

; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene

; TITLE OF INVENTION: Encoding Same

; FILE REFERENCE: 019049-000200US

; CURRENT APPLICATION NUMBER: US/09/394,630

; CURRENT FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: US 60/100,172

; PRIOR FILING DATE: 1998-09-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:amino acid

; OTHER INFORMATION: sequence obtained from protein sequencing

US-09-394-630-9

Query Match 9.0%; Score 6; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 NEIFSR 20

Db 1 NEIFSR 6

RESULT 10

US-09-394-630-14

; Sequence 14, Application US/09394630

; Patent No. 6395306

; GENERAL INFORMATION:

; APPLICANT: Cui, Xiangmin

; APPLICANT: Lu, Yuefeng

; APPLICANT: Pan Pacific Pharmaceutical, Inc.

; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene

; TITLE OF INVENTION: Encoding Same

; FILE REFERENCE: 019049-000200US

; CURRENT APPLICATION NUMBER: US/09/394,630

; CURRENT FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: US 60/100,172

; PRIOR FILING DATE: 1998-09-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:residues 55-60

; OTHER INFORMATION: of SEQ ID NO:2 obtained through protein sequencing

US-09-394-630-14

Query Match 9.0%; Score 6; DB 4; Length 6;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PNVVPK 35

Db 1 PNVVPK 6

RESULT 11

US-09-394-630-19

; Sequence 19, Application US/09394630

; Patent No. 6395306

; GENERAL INFORMATION:

; APPLICANT: Cui, Xiangmin

; APPLICANT: Lu, Yuefeng

; APPLICANT: Pan Pacific Pharmaceutical, Inc.

; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene

; TITLE OF INVENTION: Encoding Same

; FILE REFERENCE: 019049-000200US

; CURRENT APPLICATION NUMBER: US/09/394,630

; CURRENT FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: US 60/100,172

; PRIOR FILING DATE: 1998-09-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 8

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:peptide

; OTHER INFORMATION: fragment #88 from in-gel trypsin digestion

; NAME/KEY: MOD_RES

; LOCATION: (1)..(2)

; OTHER INFORMATION: xaa = unsure amino acid

US-09-394-630-19

Query Match 9.0%; Score 6; DB 4; Length 8;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 PNVVPK 35

Db 3 PNVVPK 8

RESULT 12

US-09-394-630-22

; Sequence 22, Application US/09394630

; Patent No. 6395306

; GENERAL INFORMATION:

; APPLICANT: Cui, Xiangmin

; APPLICANT: Lu, Yuefeng

; APPLICANT: Pan Pacific Pharmaceutical, Inc.

; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene

; TITLE OF INVENTION: Encoding Same

; FILE REFERENCE: 019049-000200US

; CURRENT APPLICATION NUMBER: US/09/394,630

; CURRENT FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: US 60/100,172

; PRIOR FILING DATE: 1998-09-14

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:N-terminal

; OTHER INFORMATION: sequence from Puri-#3

US-09-394-630-22

Query Match 9.0%; Score 6; DB 4; Length 10;

Best Local Similarity 100.0%; Pred. No. 7.9;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGL 6

Db 5 GGFGL 10

RESULT 13

US-09-026-633-2
 ; Sequence 2, Application US/09026633
 ; Patent No. 6025328
 ; GENERAL INFORMATION:
 ; APPLICANT: McMorris, Trevor C.
 ; APPLICANT: Kelnner, Michael J.
 ; TITLE OF INVENTION: Antitumor agents
 ; FILE REFERENCE: 103.008US1
 ; CURRENT APPLICATION NUMBER: US/09/026.633
 ; CURRENT FILING DATE: 1998-02-20
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 5
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Amino acid sequence

US-09-026-633-2

Query Match 7.5%; Score 5; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 CDGRC 25
 Db 1 CDGRC 5

RESULT 14

US-08-842-306B-20
 ; Sequence 20, Application US/08842306B
 ; Patent No. 6271197
 ; GENERAL INFORMATION:
 ; APPLICANT: Berlin, Vivian
 ; Levin, David
 ; Ohya, Yoshikazu
 ; Damagnez, Veronique
 ; Smith, Susan

TITLE OF INVENTION: ASSAYS AND REAGENTS FOR IDENTIFYING ANTI-FUNGAL AGENTS, AND USES RELATED THERETO

NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPad

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/842.306B
 FILING DATE: 23-Apr-1997

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/771,212
 FILING DATE: 20-DEC-1996
 APPLICATION NUMBER: US 08/631,319
 FILING DATE: 11-APR-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-074.04
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids

TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 20:
 US-08-842-306B-20

Query Match 7.5%; Score 5; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGG 5
 Db 1 GGFGG 5

RESULT 15

US-08-838-973B-18
 ; Sequence 18, Application US/08838973B
 ; Patent No. 6277564
 ; GENERAL INFORMATION:
 ; APPLICANT: Berlin, Vivian
 ; Damagnez, Veronique
 ; Smith, Susan

TITLE OF INVENTION: ASSAYS AND REAGENTS FOR IDENTIFYING ANTI-FUNGAL AGENTS, AND USES RELATED THERETO

NUMBER OF SEQUENCES: 48
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/838.973B
 FILING DATE: 23-Apr-1997

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/631,319
 FILING DATE: 10-APR-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-074.05
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-08-838-973B-18

Query Match 7.5%; Score 5; DB 3; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGG 5
 Db 1 GGFGG 5

Search completed: April 1, 2004, 10:07:44
 Job time : 23 secs

us-09-506-978-1.rai

Thu Apr 1 15:46:20 2004

sequence 675, App
sequence 61, Appl
sequence 59, Appl
sequence 3, Appl
sequence 394, App
sequence 938, App
sequence 1000, App
sequence 1358, App
sequence 1360, App
sequence 1376, App
sequence 1378, App
sequence 1380, App
sequence 2056, App
sequence 2058, App
sequence 2276, App
sequence 3358, App
sequence 3360, App
sequence 3478, App
sequence 3490, App
sequence 3492, App
sequence 3494, App
sequence 3528, App
sequence 3532, App
sequence 3536, App
sequence 3570, App
sequence 3572, App
sequence 3574, App
sequence 3578, App
sequence 3580, App
sequence 3596, App

ALIGNMENTS

RESULT 1

```

US-10-174-151-1
; Sequence 1, Application US/10174151
; Publication No. US2003016551A1
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM
; FILE REFERENCE: 19519-001
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/500
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-174-151-1

```

Query Match	100.0%	Score 67;	DB 14;	Length 67;
Best Local Similarity	100.0%	Pred. No. 2e-60;		
Matches	67;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

1 GGFGIGGRGKPSNEIFSRCDGRQRFQPNVWPFLCIKICAPGCVRLGLRNKKVC 60

67 00000000

RESIT.T 2

US-10-204-145-1
; Sequence 1, Application US/10204145
; Publication No. US20040023291A1
; GENERAL INFORMATION:

ppred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	67	100.0	67	14	US-10-174-151-1	Sequence 1, Appli
2	67	100.0	67	16	US-10-204-145-1	Sequence 1, Appli
3	6	9.0	10	10	US-09-572-404B-910	Sequence 910, App
4	6	9.0	10	10	US-09-572-404B-960	Sequence 960, App
5	6	9.0	10	10	US-09-572-404B-2052	Sequence 2052, Ap
6	6	9.0	10	13	US-10-044-359-26	Sequence 26, Appl
7	6	9.0	46	12	US-10-424-599-236087	Sequence 236087, A
8	6	9.0	49	14	US-10-029-386-28031	Sequence 28031, A
9	6	9.0	53	9	US-09-925-300-1261	Sequence 1261, Ap
10	6	9.0	59	9	US-09-796-652-1116	Sequence 1116, Ap
11	6	9.0	59	14	US-10-040-862-1116	Sequence 1116, Ap
12	6	9.0	59	15	US-10-057-475B-1116	Sequence 1116, Ap
13	6	9.0	59	15	US-10-454-884B-1116	Sequence 1116, Ap
14	6	9.0	60	12	US-10-424-599-235880	Sequence 235880,
15	6	9.0	65	9	US-09-764-847-675	Sequence 675, App

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:06:25 ; Search time 40 Seconds
(without alignments)
439.387 Million cell un

Title: US-09-506-978-1
Perfect score: 67
Sequence: 1 GGGGGLGGRGKCPNSNEFSR.....CRGLYLRNKKKVCVPSKCG 67

Scoring table: OLIGO

Searched: 1069545 seqs, 262320428 residues

word size :

Total number of hits satisfying chosen parameters: 349428

```
Minimum DB seq length: 0
Maximum DB seq length: 67
```

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

/cgm2_6/ptodata/2/pubpaa/us07_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/FACTUS_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp;
 /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp;

ppred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```

; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-204-145-1

Query Match          100.0%; Score 67; DB 16; Length 67;
Best Local Similarity 100.0%; Pred. No. 2e-60;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRGKCPNIEIFSRDGRQCFPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
DB 1 GGFGLGGRGKCPNIEIFSRDGRQCFPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60

QY 61 VPRSKCG 67
DB 61 VPRSKCG 67

RESULT 3
US-09-572-404B-910
; Sequence 910, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 910
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in KRT9 at 473-482 and may interact with Sequen
; OTHER INFORMATION: in this patent.
US-09-572-404B-910

Query Match          9.0%; Score 6; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GLGGRG 10
DB 1 GLGGRG 6

RESULT 4
US-09-572-404B-960
; Sequence 960, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 960
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Androctonus mauretanicus
US-10-044-359-26

Query Match          9.0%; Score 6; DB 13; Length 35;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GLGGRG 10
DB 1 GLGGRG 6

RESULT 5
US-09-572-404B-2052
; Sequence 2052, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 2052
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in KRT9 at 473-482 and may interact with Sequen
; OTHER INFORMATION: in this patent.
US-09-572-404B-2052

Query Match          9.0%; Score 6; DB 10; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GLGGRG 10
DB 1 GLGGRG 6

RESULT 6
US-10-044-359-26
; Sequence 26, Application US/10044359
; Publication No. US20020160454A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, Rafael
; APPLICANT: Wong, James F.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: SCORPION TOXINS
; FILE REFERENCE: BB1367 US NA
; CURRENT APPLICATION NUMBER: US/10/044,359
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/599,416
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/140,227
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Androctonus mauretanicus
US-10-044-359-26

Query Match          9.0%; Score 6; DB 13; Length 35;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GLGGRG 10
DB 1 GLGGRG 6
```

QY 7 GGRGKC 12
| | | | |
Db 20 GGRGKC 25

RESULT 7
US-10-424-599-236087
; Sequence 236087, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 236087
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55214C.1.pep
US-10-424-599-236087

Query Match 9.0%; Score 6; DB 12; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GLGGRG 10
| | | | |
Db 22 GLGGRG 27

RESULT 8
US-10-029-386-28031
; Sequence 28031, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28031
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR22 76.0
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
US-10-029-386-28031

Query Match 9.0%; Score 6; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 DGRQCR 27

Db 8 DGRQCR 13
| | | | |

RESULT 9
US-09-925-300-1261
; Sequence 1261, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1261
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1261

Query Match 9.0%; Score 6; DB 9; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGLGGR 9
| | | | |
Db 46 GGLGGR 51

RESULT 10
US-09-796-692-1116
; Sequence 1116, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1116
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-692-1116

Query Match 9.0%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGGL 6
Db 27 GGFGGL 32

RESULT 11
US-10-040-862-1116
Sequence 1116, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1116
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-10-040-862-1116

Query Match 9.0%; Score 6; DB 14; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGGL 6
Db 27 GGFGGL 32

RESULT 12
US-10-057-475B-1116
Sequence 1116, Application US/10057475B
Publication No. US2004002068A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1116
LENGTH: 59
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-475B-1116

Query Match 9.0%; Score 6; DB 15; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGGL 6
Db 27 GGFGGL 32

RESULT 13
US-10-154-884B-1116
Sequence 1116, Application US/10154884B
Publication No. US20040005561A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc W.
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
TITLE OF INVENTION: Hematological Malignancies
FILE REFERENCE: 014058-013521US
CURRENT APPLICATION NUMBER: US/10/154,884B

```

; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/219,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1116
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-154-884B-1116

Query Match          9.0%; Score 6; DB 15; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGGL 6
Db 27 GFGGGL 32

RESULT 14
US-10-424-599-235880
; Sequence 235880, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 235880
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(60)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT3847_55027C.1.pep
; US-10-424-599-235880

Query Match          9.0%; Score 6; DB 12; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GFGGLG 7
Db 44 GFGGLG 49
```

```

RESULT 15
US-09-764-847-675
; Sequence 675, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 675
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (14)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-847-675

Query Match          9.0%; Score 6; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGGRGK 11
Db 45 LGGRGK 50

Search completed: April 1, 2004, 10:12:14
Job time : 41 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:02:54 ; Search time 20 Seconds
(without alignments)
322.242 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 67
Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRGLYLRNKKKVCVPRSKCG 67

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 19407

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	6	9.0	35	1 NTSRPM	neurotoxin P2 - sc
2	5	7.5	27	2 A31508	34K nucleolar prot
3	5	7.5	30	2 C45087	cysteine proteins
4	5	7.5	31	2 A23887	fibrillarin - rat
5	5	7.5	32	2 H81215	hypothetical prote
6	5	7.5	33	2 A42592	precorrin-6x reduc
7	5	7.5	35	2 A61375	basic fibroblast g
8	5	7.5	36	1 A48850	chloride channel l
9	5	7.5	36	2 JN0402	insect toxin 1l -
10	5	7.5	40	2 S52342	hypothetical prote
11	5	7.5	42	2 A71429	hypothetical prote
12	5	7.5	43	2 S41388	protein 3a - human
13	5	7.5	47	2 I79424	MHC HLA-DR-beta-1
14	5	7.5	48	2 S22398	tral protein - Esc
15	5	7.5	51	2 T09005	variant homolog s
16	5	7.5	53	2 AG0053	probable exported
17	5	7.5	54	2 C91282	hypothetical prote
18	5	7.5	54	2 E86123	hypothetical prote
19	5	7.5	56	2 S33789	ribosomal protein
20	5	7.5	58	2 F84303	50S ribosomal prot
21	5	7.5	58	2 H89964	hypothetical prote
22	5	7.5	59	2 A86137	hypothetical prote
23	5	7.5	59	2 AF1072	hypothetical prote
24	5	7.5	61	2 D97051	hypothetical prote
25	5	7.5	65	2 A81900	hypothetical prote
26	5	7.5	67	2 F75128	ssu ribosomal prot
27	4	6.0	4	2 A53284	T-cell receptor be
28	4	6.0	5	2 B22565	R-phycoerythrin al
29	4	6.0	6	2 I51434	H4 histone - Afric

30	4	6.0	6	2	PT0727	T-cell receptor be
31	4	6.0	7	2	PT0663	T-cell receptor be
32	4	6.0	8	2	PT0725	T-cell receptor be
33	4	6.0	9	2	PT0670	T-cell receptor be
34	4	6.0	10	2	F60527	sperm-activating p
35	4	6.0	10	2	S06964	hypothetical prote
36	4	6.0	10	2	PT0632	T-cell receptor be
37	4	6.0	10	2	PT0664	T-cell receptor be
38	4	6.0	13	2	PS0443	potassium channel
39	4	6.0	13	2	PH1596	Ig H chain V-D-J r
40	4	6.0	14	2	I51432	histone H4-1 precu
41	4	6.0	14	2	A17150	glucose 1-dehydrog
42	4	6.0	15	2	A27504	histone H2A - mous
43	4	6.0	15	2	PH1616	Ig H chain V-D-J r
44	4	6.0	16	2	A11488	taurocyamine kinas
45	4	6.0	16	2	PT0296	Ig heavy chain CDR

ALIGNMENTS

RESULT 1

NTSRPM
neurotoxin P2 - scorpion (Androctonus mauretanicus)
C:Species: Androctonus mauretanicus mauretanicus
C:Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 23-Aug-1996
C:Accession: A01758
R:Rosso, J.P.; Rochat, H.
Toxinon 23, 113-125, 1985
A:Title: Characterization of ten proteins from the venom of the Moroccan scorpion Androc
A:Reference number: A94318; MUID:85193276; PMID:3992595
A:Accession: A01758
A:Molecule type: protein
A:Residues: 1-35 <ROS>
C:Superfamily: scorpion neurotoxin
C:Keywords: neurotoxin; venom
P:1-18,4-25,15-30,19-32/Disulfide bonds: #status predicted

Query Match 9.0%; Score 6; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	7	GGRGKC	12
DB	20	GGRGKC	25

RESULT 2

A31508
34K nucleolar protein B-36 - slime mold (Physarum polycephalum) (fragment)
C:Species: Physarum polycephalum
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 14-Aug-1998
C:Accession: A31508
R:Christensen, M.E.; Fuxa, K.P.
Biochem. Biophys. Res. Commun. 155, 1278-1283, 1988
A:Title: The nucleolar protein, B-36, contains a glycine and dimethylarginine-rich sequ
A:Reference number: A31508; MUID:89025797; PMID:3140806
A:Accession: A31508
A:Molecule type: protein
A:Residues: 1-27 <CHR>
C:Superfamily: nucleolin; ribonucleoprotein repeat homology
C:Keywords: methylated amino acid; nucleolus
F:5,11,16,19/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status experimental

Query Match 7.5%; Score 5; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGFGG	5
DB	6	GGFGG	10

RESULT 3

G45087
cysteine proteinase homolog (clone PCR31) - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C/Accession: G45087
R;Petanceska, S.; Devi, L.
J. Biol. Chem. 267, 26038-26043, 1992
A/Title: Sequence analysis, tissue distribution, and expression of rat cathepsin S.
A/Reference number: A45087; MUID:93100327; PMID:1281481
A/Accession: G45087
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-30 <PEP>
A/Experimental source: brain
A/Note: sequence extracted from NCBI backbone (NCBIP:123684)

Query Match 7.5%; Score 5; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 LGGRG 10
|||||
DB 16 LGGRG 20

RESULT 4

A23887
fibrillarin - rat (fragment)
N/Alternate names: nucleolar scleroderma 34K antigen
C/Species: Rattus norvegicus (Norway rat)
C/Date: 28-Jul-1987 #sequence_revision 28-Jul-1987 #text_change 18-Sep-1998
C/Accession: A23887
R;Lischwe, M.A.; Ochs, R.L.; Reddy, R.; Cook, R.G.; Yeoman, L.C.; Tan, E.M.; Reichlin, M.
J. Biol. Chem. 260, 14304-14310, 1985
A/Title: Purification and partial characterization of a nucleolar scleroderma antigen (M
A/Reference number: A23887; MUID:86033920; PMID:2414294
A/Accession: A23887

A/Molecule type: protein
A/Residues: 1-31 <UIS>
A/Experimental source: Novikoff hepatoma ascites cells (rat)
C/Superfamily: human fibrillarin
C/Keywords: methylated amino acid; nucleolus
F,8,15,21,24,28,31/Modified site: omega-N,omega-N-dimethylarginine (Arg) #status experim
Query Match 7.5%; Score 5; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGFGG 5
|||||
DB 10 GGFGG 14

RESULT 5

H81215
hypothetical protein NMB0297 [imported] - Neisseria meningitidis (strain MC58 serogroup
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C/Accession: H81215
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, C.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A/Reference number: A81000; MUID:20175755; PMID:10710307
A/Accession: H81215
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-32 <PEP>
A/Cross-references: GB:AE002386; GB:AE002098; NID:g7225512; PIDN:AAF40748.1; PID:g722552
A/Experimental source: serogroup B, strain MC58

C/Genetics:

A/Gene: NMB0297

Query Match 7.5%; Score 5; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GGRGK 11
|||||
DB 20 GGRGK 24

RESULT 6

A42592
precorrin-6x reductase (EC 1.-.-.-) - Pseudomonas sp. (strain SC510) (fragments)
C/Species: Pseudomonas sp.
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Mar-1996
C/Accession: A42592
R;Blanchet, F.; Thibaut, D.; Famechon, A.; Debussche, L.; Cameron, B.; Crouzet, J.
J. Bacteriol. 174, 1036-1042, 1992
A/Title: Precorrin-6x reductase from Pseudomonas denitrificans: purification and charac
A/Reference number: A42592; MUID:92121090; PMID:1732193
A/Accession: A42592
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-33 <BLA>
C/Keywords: oxidoreductase

Query Match 7.5%; Score 5; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGG 5
|||||
DB 8 GGFGG 12

RESULT 7

A61375
basic fibroblast growth factor - rat (fragment)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C/Accession: A61375
R;Powell, P.P.; Klagsbrun, M.
J. Cell. Physiol. 148, 203-210, 1991
A/Title: Three forms of rat basic fibroblast growth factor are made from a single mRNA
A/Reference number: A61375; MUID:91349212; PMID:1880150
A/Accession: A61375
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-35 <POW>
C/Genetics:
A/Start codon: AGG
C/Keywords: alternative initiators

Query Match 7.5%; Score 5; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGGRG 10
|||||
DB 9 LGGRG 13

RESULT 8

A48850
chloride channel ligand chlorotoxin - Egyptian scorpion
C/Species: Leiurus quinquestriatus (Egyptian scorpion)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: A48850
R;DeBin, J.A.; Maggio, J.E.; Strichartz, G.R.
Am. J. Physiol. 264, C361-C369, 1993
A/Title: Purification and characterization of chlorotoxin, a chloride channel ligand fr

A;Reference number: A48850; MUID:93191044; PMID:8383429
A;Accession: A48850
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-36 <DEB>
A;Experimental source: venom
A;Note: sequence extracted from NCBI backbone (NCBIP:126513)
C;Superfamily: scorpion neurotoxin
C;Keywords: chloride channel inhibitor
F;2-19,5-28,16-33,20-35/Disulfide bonds: #status predicted

Query Match 7.5%; Score 5; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRGKC 12
|||||
DB 24 GRGKC 28

RESULT 9
JN0402
insect toxin II - lesser Asian scorpion
C;Species: Mesobuthus eupeus (lesser Asian scorpion)
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 31-Dec-1993
C;Accession: JN0402
R;Zhdanova, L.N.; Adamovich, T.B.; Nazimov, I.V.; Grishin, E.V.; Ovchinnikov, Y.A.
Bioorg. Khim. 3, 485-493, 1977
A;Title: Amino acid sequence of insectotoxin II from the venom of middle-Asian scorpion
A;Reference number: JN0402
A;Accession: JN0402
A;Molecule type: protein
A;Residues: 1-36 <ZHD>
A;Note: all cysteine residues are involved in intrachain disulfide bonds
C;Superfamily: scorpion neurotoxin
C;Keywords: neurotoxin; venom

Query Match 7.5%; Score 5; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRGKC 12
|||||
DB 22 GRGKC 26

RESULT 10
S52342
hypothetical protein - Lactococcus lactis (fragment)
C;Species: Lactococcus lactis
C;Date: 08-May-1995 #sequence_revision 06-Dec-1996 #text_change 15-Oct-1999
R;Waterfield, N.R.; LePage, R.W.; Wilson, P.W.; Wells, J.M.
submitted to the EMBL Data Library, February 1995
A;Description: The isolation of lactococcal promoters and their use to investigate bacteriophage promoters
A;Reference number: S52330
A;Accession: S52342
A;Molecule type: DNA
A;Residues: 1-40 <WAT>
A;Cross-references: EMBL:Z48200; NID:g666057; PIDN:CAA88236.1; PID:g666058
C;Genetics:
A;Start codon: TTG

Query Match 7.5%; Score 5; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 NKKKV 59
|||||
DB 2 NKKKV 6

RESULT 11

A71429
hypothetical protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: Columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C;Accession: A71429
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirr, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel, van der, T.; Hempel, S.; Kottler, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech, E.; Chaiwatiz, N.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: A71429
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-42 <BEV>
A;Cross-references: GB:Z97340; NID:g2244950; PID:e327498; PID:g2244982
C;Genetics:
A;Map position: 4COP9-4G3845

Query Match 7.5%; Score 5; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGG 5
|||||
DB 26 GGFGG 30

RESULT 12
S41388
protein 3a - human adenovirus 3 (fragment)
C;Species: Mastadenovirus h3 (human adenovirus 3)
C;Date: 06-Jan-1995 #sequence_revision 30-Jan-1998 #text_change 26-Aug-1999
C;Accession: S41388
R;Curange, A.; Chroboczek, J.; Jacrot, B.
submitted to the EMBL Data Library, January 1994
A;Description: The penton base of human adenovirus type 3 has the RGD motif.
A;Reference number: S41388
A;Accession: S41388
A;Molecule type: DNA
A;Residues: 1-43 <CUZ>
A;Cross-references: EMBL:Z29487; NID:g444048; PIDN:CAA82621.1; PID:g444049
A;Experimental source: serotype 3
C;Superfamily: adenovirus penton base associated protein

Query Match 7.5%; Score 5; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGGRG 10
|||||
DB 23 LGGRG 27

RESULT 13
I79424
MHC HLA-DR-beta-1 chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 23-Jul-1999
R;Gregersen, P.K.; Shen, M.; Song, Q.
Proc. Natl. Acad. Sci. U.S.A. 83, 2642-2646, 1986
A;Title: Molecular diversity of HLA-DR4 haplotypes.
A;Reference number: I59062; MUID:86206008; PMID:3458223
A;Accession: I79424
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-47 <RES>
A;Cross-references: GB:M15073; NID:g188370; PIDN:AAA59814.1; PID:g386941

Search completed: April 1, 2004, 10:06:19
Job time : 22 secs

C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 7.5%; Score 5; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGGRG 10
Db 33 LGGRG 37

RESULT 14

S22998
traL protein - Escherichia coli plasmid R751 (fragment)
C:Species: Escherichia coli
C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 08-Oct-1999
C:Accession: S22998
R:Zieglerin, G.; Pansegrau, W.; Strack, B.; Balzer, D.; Kroeger, M.; Kruft, V.; Lanka, E.
DNA Seq. 1, 303-327, 1991
A:Title: Nucleotide sequence and organization of genes flanking the transfer origin of P
A:Reference number: S22992; MUID:92190548; PMID:1665997
A:Accession: S22998
A:Molecule type: DNA
A:Residues: 1-48 <ZIE>
A:Cross-references: EMBL:X54458; NID:942656; PIDN:CAA38333.1; PID:942663
C:Genetics:
A:Gene: traL
A:Genome: plasmid

Query Match 7.5%; Score 5; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 KPLCI 39
Db 33 KPLCI 37

RESULT 15

T09005
variacin homolog scnA1 precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T09005
R:Hynes, W.L.; Ferretti, J.J.; Tagg, J.R.
Appl. Environ. Microbiol. 59, 1969-1971, 1993
A:Title: Cloning of the gene encoding streptococin A-PF22, a novel lantibiotic produced
A:Reference number: A58598; MUID:93319301; PMID:832813
A:Accession: T09005
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-51 <HYN>
A:Cross-references: EMBL:AF026542; NID:92502065; PIDN:AAB92601.1; PID:92502069
C:Genetics:
A:Gene: scnA1
C:Superfamily: unassigned lanthionine-containing peptides
C:Keywords: antibiotic; lanthionine
F:1-27/Domain: propeptide #status predicted <SIG>
F:28-51/Product: variacin homolog #status predicted <MAT>
F:34-39/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys) #status predicted
F:36-50/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status predicted
F:43-51/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status predicted
F:49/Modified site: (Z)-dehydrobutyrine (Thr) #status predicted

Query Match 7.5%; Score 5; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGGRG 10
Db 25 LGGRG 29

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:00:49 ; Search time 13 Seconds
(without alignments)
268.362 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 67

Sequence: 1 GGFGLGGRKCPSEIFSR.....CRGLYLRNKKVCVPSKCG 67

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7280

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	9.0	35	1	SXCP ANDMA
2	5	7.5	14	1	SAP2_ARBPU
3	5	7.5	27	1	FBRU_PHYPO
4	5	7.5	30	1	ITR4_CUCSA
5	5	7.5	31	1	FBRU_RAT
6	5	7.5	36	1	SCX1_BUTEU
7	5	7.5	35	1	SCXL_LEIQU
8	5	7.5	48	1	DHSU_CHLIT
9	5	7.5	56	1	RL37_HALMA
10	5	7.5	58	1	RL37_HALN1
11	5	7.5	67	1	R17E_PYRAB
12	5	7.5	67	1	R17E_PYRFU
13	5	7.5	67	1	R17E_PYRHO
14	4	6.0	12	1	CD14_LITXA
15	4	6.0	12	1	UKA2_HUMAN
16	4	6.0	15	1	CDN2_LITGI
17	4	6.0	15	1	CDN3_LITGI
18	4	6.0	15	1	CDN5_LITCE
19	4	6.0	15	1	CDN6_LITCE
20	4	6.0	16	1	KTRC_AREMA
21	4	6.0	17	1	LCK_RAT
22	4	6.0	20	1	KORC_METTM
23	4	6.0	23	1	CP23_SPOER
24	4	6.0	23	1	PAP1_HELVI
25	4	6.0	23	1	PAP1_MANSE
26	4	6.0	23	1	PAP1_SPOEX
27	4	6.0	23	1	PAP2_HELVI
28	4	6.0	23	1	PAP2_MANSE
29	4	6.0	23	1	PAP2_SPOEX
30	4	6.0	23	1	PAP3_SPOEX
31	4	6.0	24	1	LPER_STEPR
32	4	6.0	24	1	PBRW_WHEAT
33	4	6.0	25	1	ANDT_ANDAU

34	4	6.0	25	1	CXOB_CONMA
35	4	6.0	25	1	GBP_APAKA
36	4	6.0	26	1	CXOC_CONCT
37	4	6.0	26	1	RL36_DESDE
38	4	6.0	29	1	CXOC_CONMA
39	4	6.0	29	1	ITR2_BRYDI
40	4	6.0	29	1	KDPF_ECOLI
41	4	6.0	29	1	TAT_HV123
42	4	6.0	30	1	CX7A_CONTU
43	4	6.0	30	1	ITR2_ECBEL
44	4	6.0	30	1	TX2_HETVE
45	4	6.0	30	1	TX2_THRPR

ALIGNMENTS

RESULT 1

ID	SCXP ANDMA	STANDARD;	PRT;	35 AA.
AC	P01498;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neurotoxin P2.			
OS	Androctonus mauretanicus mauretanicus (Scorpion).			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;			
OC	Euthioidea; Buthidae; Androctonus.			
OX	NCBI_TaxID=6860;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Venom;			
RX	MEDLINE=85193276; PubMed=392595;			
RA	Rosso J.P., Rochat H.;			
RT	"Characterization of ten proteins from the venom of the Moroccan scorpion Androctonus mauretanicus mauretanicus, six of which are toxic to the mouse."			
RT	Toxin; Neurotoxin.			
RL	Toxicon 23:113-125(1985).			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Expressed by the venom gland.			
CC	-!- SIMILARITY: Belongs to the short scorpion toxin family.			
DR	PIR; A01758; NTSRPM.			
DR	HSSP; P15222; LSIS.			
DR	InterPro; IPR007958; toxin_5.			
DR	Pfam; PF05294; toxin_5; 1.			
KW	Toxin; Neurotoxin.			
FT	DISULFID 1 18 BY SIMILARITY.			
FT	DISULFID 4 25 BY SIMILARITY.			
FT	DISULFID 15 30 BY SIMILARITY.			
FT	DISULFID 19 32 BY SIMILARITY.			
SQ	SEQUENCE 35 AA; 3673 MW; 213E8926289B5A CRC64;			

Query Match 9.0%; Score 6; DB 1; Length 35;

Best Local Similarity 100.0%; Pred. No. 8.7; 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0;

OY 7 GGRGKC 12

DB 20 GGRGKC 25

RESULT 2

ID	SAP2_ARBPU	STANDARD;	PRT;	14 AA.
AC	P11760;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Resact (Sperm-activating peptide) (SAP-IIA).			
OS	Arbacia punctulata (Punctate sea urchin).			
OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Arbacia.			
OC	Echinoidea; Echinoidea; Echinacea; Arbacoidea; Arbacoidea; Arbacoidea.			
OX	NCBI_TaxID=7641;			

```

RN RP SEQUENCE.
RC TISSUE=Egg;
RX MEDLINE=85054381; PubMed=6150045;
RA Suzuki N., Shimomura H., Radany E.W., Ramarao C.S., Ward G.E.,
RA Bentley J.K., Garbers D.L.;
RT "A peptide associated with eggs causes a mobility shift in a major
RT plasma membrane protein of spermatozoa.";
RL J. Biol. Chem. 259:14874-14879(1984).
RN [2]
RP DISULFIDE BOND.
RX MEDLINE=92097763; PubMed=1756858;
RA Yoshino K.-I., Takao T., Shimonishi Y., Suzuki N.;
RT "Determination of the amino acid sequence of an intramolecular
RT disulfide linkage-containing sperm-activating peptide by tandem mass
RT spectrometry.";
RL FEBS Lett. 294:179-182(1991).
CC -!- FUNCTION: Cause stimulation of sperm respiration and motility
CC through intracellular alkalization, transient elevations of
CC cAMP, cGMP and calcium levels in sperm cells, and transient
CC activation and subsequent inactivation of the membrane form of
CC guanylate cyclase.
CC -!- SIMILARITY: SMALL TO S.PURPURATUS SPERACT.
KW Amidation.
FT MOD_RES 1 8
FT DISULFID 14 14
FT SEQUENCE 14 AA; 1246 MW; 39745AA33BEE41B8 CRC64;
SQ SEQUENCE 14 AA; 1246 MW; 39745AA33BEE41B8 CRC64;

Query Match 7.5%; Score 5; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 APGCV 47
Db 5 APGCV 9

RESULT 3
ID FBRL PHYPO STANDARD; PRT; 27 AA.
AC P22508;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrillar (34 kDa nucleolar protein B-36) (Fragment).
OS Physarum polycephalum (Slime mold).
OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physariida;
OC Physarum.
OC NCBI_TaxID=5791;
RN [1]
RP SEQUENCE.
RX MEDLINE=89025797; PubMed=3140806;
RA Christensen M.E., Fuxa K.P.;
RT "The nucleolar protein, B-36, contains a glycine and
RT dimethylarginine-rich sequence conserved in several other nuclear
RT RNA-binding proteins.";
RL Biochem. Biophys. Res. Commun. 155:1278-1283(1988).
CC -!- FUNCTION: Fibrillar is a component of a nucleolar small nuclear
CC ribonucleoprotein particle thought to participate in the first
CC step in processing preribosomal RNA. It is associated with the U3,
CC U8 and U13 small nuclear RNAs.
CC -!- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus.
CC -!- SIMILARITY: Belongs to the fibrillar family.
DR PIR: A31508; A31508.
DR InterPro: IPR000692; Fibrillar.
DR PROSITE: PS00566; FIBRILLARIN; PARTIAL.
KW Nuclear protein; Methylation; Ribonucleoprotein; rRNA processing;
KW RNA-binding.
FT MOD_RES 5 5
FT MOD_RES 11 11
FT MOD_RES 16 16
FT MOD_RES 19 19
FT NON_TER 27 27

SQ SEQUENCE 27 AA; 2462 MW; F76AD7F8FAF442DA CRC64;

Query Match 7.5%; Score 5; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGG 5
Db 6 GFGGG 10

RESULT 4
ID FBRL PHYPO STANDARD; PRT; 30 AA.
AC P10232;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trypsin inhibitor IV (CSTI-IV).
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OC NCBI_TaxID=3659;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=85149300; PubMed=3977882;
RA Weczkorek M., Otlewski J., Cook J., Parks K., Ielux J.,
RA Willnowska-Pelc A., Polanowski A., Wilusz T., Laskowski M. Jr.;
RT "The squash family of serine proteinase inhibitors. Amino acid
RT sequences and association equilibrium constants of inhibitors from
RT squash, summer squash, zucchini, and cucumber seeds.";
RL Biochem. Biophys. Res. Commun. 126:646-652(1985).
CC -!- SIMILARITY: Belongs to the squash-type serine protease inhibitor
CC family.
DR HSP: P01074; ICTI.
DR InterPro: IPR000737; Squash.
DR Pfam: PF00299; squash; 1.
DR PRINTS: PR00293; SQUASHINHBTR.
DR ProDom: PD003401; Squash; 1.
DR SMART: SM00286; PTI; 1.
DR PROSITE: PS00286; SQUASH_INHIBITOR; 1.
KW Serine protease inhibitor.
FT ACT_SITE 5 6
FT ACT_SITE 3 20
FT DISULFID 10 22
FT DISULFID 16 29
FT SEQUENCE 30 AA; 3429 MW; 7A4B870E7D2A088F CRC64;

Query Match 7.5%; Score 5; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 PGCV 48
Db 18 PGCV 22

RESULT 5
ID FBRL RAT STANDARD; PRT; 31 AA.
AC P22509;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrillar (Nucleolar protein 1) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]

```

RP SEQUENCE.
 RX MEDLINE=86033920; PubMed=2414294;
 RA Liesche M.A., Ochs R.L., Reddy R., Cook R.G., Yeoman L.C., Tan E.M.,
 RA Reichlin M., Busch H.;
 RT "Purification and partial characterization of a nucleolar scleroderma
 RL antigen (Mr = 34,000; pI, 8.5) rich in NG,NG-dimethylarginine.";
 RL J. Biol. Chem. 260:14304-14310(1985).
 RN [2]
 RP SEQUENCE OF 1-28, AND INTERACTION WITH NOLC1.
 RX MEDLINE=20143579; PubMed=10679015;
 RA Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;
 RT "Conserved composition of mammalian box H/ACA and box C/D small
 RL nucleolar ribonucleoprotein particles and their interaction with the
 RL common factor Nopp140.";
 RL Mol. Biol. Cell 11:567-577(2000).
 CC -1- FUNCTION: Fibrillarin is a component of a nucleolar small nuclear
 CC ribonucleoprotein particle thought to participate in the first
 CC step in processing preribosomal RNA. It is associated with the U3,
 CC U8 and U13 small nuclear RNAs.
 CC -1- SUBUNIT: Interacts with NOLC1/Nopp140.
 CC -1- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus.
 CC -1- SIMILARITY: Belongs to the fibrillarin family.
 DR PIR; A23887; A23887.
 DR PROSITE; PS00566; FIBRILLARIN; PARTIAL.
 DR Nuclear protein, Methylation, Ribonucleoprotein; rRNA processing;
 KW RNA-binding.
 FT MOD_RES 8 8 METHYLATION (DI-).
 FT MOD_RES 15 15 METHYLATION (DI-).
 FT MOD_RES 21 21 METHYLATION (DI-).
 FT MOD_RES 24 24 METHYLATION (DI-).
 FT MOD_RES 28 28 METHYLATION (DI-).
 FT MOD_RES 31 31 METHYLATION (DI-).
 FT CONFLICT 2 2 K -> D (IN REF. 2).
 FT NON_TER 31 31
 SQ SEQUENCE 31 AA; 2965 MW; 93EBCC102847A363 CRC64;

Query Match 7.5%; Score 5; DB 1; Length 31;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGFGG 5
 Db 10 GGFGG 14

RESULT 6
 SCX1_BUTEU STANDARD; PRT; 36 AA.
 AC P15220;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Insectotoxin II.
 OS Buthus eupeus (lesser Asian scorpion) (Mesobuthus eupeus).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Mesobuthus.
 OX NCBI_TaxID=34648;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RA Adamovich T.B., Nasimov I.V., Grishin E.V., Ovchinnikov Y.A.;
 RT "Amino acid sequence of insectotoxin II from the venom of
 RT middle-asian scorpion Buthus eupeus";
 RL Bioorg. Khim. 3:485-493(1977).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: Belongs to the short scorpion toxin family.
 DR PIR; JN0402; JN0402.
 DR HSSP; P15222; 1SIS.
 DR InterPro; IPR007958; toxin_5.
 DR Pfam; PF05294; toxin_5; 1.
 KW Toxin; Neurotoxin.

FT DISULFID 2 19 BY SIMILARITY.
 FT DISULFID 5 26 BY SIMILARITY.
 FT DISULFID 16 31 BY SIMILARITY.
 FT DISULFID 20 33 BY SIMILARITY.
 SQ SEQUENCE 36 AA; 4011 MW; C75228525076C284 CRC64;
 Query Match 7.5%; Score 5; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 8 GRGKC 12
 Db 22 GRGKC 26
 RESULT 7
 SCX1_LEIQU STANDARD; PRT; 36 AA.
 AC P45639;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chlorotoxin (CTX).
 OS Leiurus quinquestriatus quinquestriatus (Egyptian scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Leiurus.
 OX NCBI_TaxID=6885;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=93191044; PubMed=8383429;
 RA Debin J.A., Maggio J.E., Strichartz G.R.;
 RT "Purification and characterization of chlorotoxin, a chloride channel
 RL ligand from the venom of the scorpion.";
 RL Am. J. Physiol. 264:C361-C369(1993).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=95118979; PubMed=7819188;
 RA Lippens G., Najib J., Wodak S.J., Tartar A.;
 RT "NMR sequential assignments and solution structure of chlorotoxin, a
 RT small scorpion toxin that blocks chloride channels.";
 RL Biochemistry 34:13-21(1995).
 CC -1- FUNCTION: Chloride channel ligand. Blocks small-conductance
 CC chloride channels.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: Belongs to the short scorpion toxin family.
 DR PIR; A48850; A48850.
 DR PDB; 1CHL; 07-FEB-95.
 DR InterPro; IPR007958; toxin_5.
 DR Pfam; PF05294; toxin_5; 1.
 KW Toxin; Neurotoxin; Chloride channel inhibitor;
 KW Ionic channel inhibitor; 3D-structure.
 FT DISULFID 2 19
 FT DISULFID 5 28
 FT DISULFID 16 33
 FT DISULFID 20 35
 FT STRAND 2 2
 FT HELIX 13 20
 FT TURN 21 21
 FT STRAND 27 29
 FT TURN 30 31
 FT STRAND 32 34
 SQ SEQUENCE 36 AA; 4005 MW; 14A9F57559C6E92A CRC64;
 Query Match 7.5%; Score 5; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 8 GRGKC 12
 Db 24 GRGKC 28

```

RT A resolution."
RL Science 289:905-920(2000).
CC -1- SIMILARITY: Belongs to the L37E family of ribosomal proteins.
DR PIR; S33789; S33789
DR PDB; 1FKK; 26-SEP-01.
DR PDB; 1K8A; 19-JUL-02.
DR PDB; 1K9W; 19-JUL-02.
DR PDB; 1K01; 19-JUL-02.
DR PDB; 1MLK; 23-AUG-02.
DR PDB; 1M90; 06-SEP-02.
DR HAMAP; MF_00547; -, 1.
DR InterPro; IPR001569; Ribosomal_L37E.
DR Pfam; PF01907; Ribosomal_L37E; 1.
DR ProDom; PD005132; Ribosomal_L37E; 1.
DR PROSITE; PS01077; RIBOSOMAL_L37E; 1.
KW Ribosomal protein; 3D-structure.
FT INIT_MET 0
SQ SEQUENCE 56 AA; 6193 MW; EF5705E7873112EC CRC64;

Query Match 7.5%; Score 5; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 KKKVC 60
DB 30 KKKVC 34

RESULT 10
RL37 HALM1 STANDARD; PRT; 58 AA.
ID RL37 HALM1 STANDARD; PRT; 58 AA.
AC QHPF3; 2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L37e.
GN RPL37E OR VNG1494G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- SIMILARITY: Belongs to the L37E family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005064; AAG19794.1; -.
DR PIR; F84303; F84303.
DR HAMAP; MF_00547; -, 1.
DR InterPro; IPR001569; Ribosomal_L37E.
DR Pfam; PF01907; Ribosomal_L37E; 1.
DR ProDom; PD005132; Ribosomal_L37E; 1.
DR PROSITE; PS01077; RIBOSOMAL_L37E; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 58 AA; 6357 MW; E0AECA19DAD7F112 CRC64;

```

```

RT A resolution."
RL Science 289:905-920(2000).
CC -1- SIMILARITY: Belongs to the L37E family of ribosomal proteins.
DR PIR; S33789; S33789
DR PDB; 1FKK; 26-SEP-01.
DR PDB; 1K8A; 19-JUL-02.
DR PDB; 1K9W; 19-JUL-02.
DR PDB; 1K01; 19-JUL-02.
DR PDB; 1MLK; 23-AUG-02.
DR PDB; 1M90; 06-SEP-02.
DR HAMAP; MF_00547; -, 1.
DR InterPro; IPR001569; Ribosomal_L37E.
DR Pfam; PF01907; Ribosomal_L37E; 1.
DR ProDom; PD005132; Ribosomal_L37E; 1.
DR PROSITE; PS01077; RIBOSOMAL_L37E; 1.
KW Ribosomal protein; 3D-structure.
FT INIT_MET 0
SQ SEQUENCE 56 AA; 6193 MW; EF5705E7873112EC CRC64;

Query Match 7.5%; Score 5; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGFG 5
DB 10 GGGFG 14

RESULT 9
RL37 HALMA STANDARD; PRT; 56 AA.
ID RL37 HALMA STANDARD; PRT; 56 AA.
AC P32410;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L37e (L35e).
GN RPL37E.
OS Haloarcula marismortui (Halobacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
CX NCBI_TaxID=2238;
RN [1]
RP SEQUENCE.
RX MEDLINE=93277953; PubMed=8504167;
RA Bergmann U., Wittmann-Liebold B.;
RT "H35e and HLA: primary structure of two very basic and cysteine-rich
RT ribosomal proteins from Haloarcula marismortui."
RL Biochim. Biophys. Acta 1173:195-200(1993).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC STRAIN=ATCC 43049;
RX MEDLINE=20396344; PubMed=10937989;
RA Ban N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;
RT "The complete atomic structure of the large ribosomal subunit at 2.4

```

```

Query Match          7.5%; Score 5; DB 1; Length 59;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

QY 56 KKKVC 60
DB 31 KKKVC 35

RESULT 11
R17E_PYRAB
ID R17E_PYRAB STANDARD; PRT; 67 AA.
AC QSV0G0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 30S ribosomal protein S17e.
GN RPS17E OR PYRAB08290 OR PAB7207.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GE5 / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi.";
RL Mol. Microbiol. 47:1495-1512(2003).
CC -!- SIMILARITY: Belongs to the S17E family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AJ248285; CAB49743.1; -.
CC HAMAP; MF 00511; -. 1.
CC InterPro; IPR001210; Ribosomal_S17E.
CC Pfam; PF00833; Ribosomal_S17e; 1.
CC PROSITE; PS00712; RIBOSOMAL_S17E; 1.
CC Ribosomal protein; Complete proteome.
SQ SEQUENCE 67 AA; 8040 MW; BB59539F9835AC6 CRC64;

Query Match          7.5%; Score 5; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 NKXKV 59
DB 31 NKXKV 35

RESULT 13
R17E_PYRHO
ID R17E_PYRHO STANDARD; PRT; 67 AA.
AC P58503;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S17e.
GN RPS17E OR PH1316.1.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kavarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Furahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
RN [2]
RP IDENTIFICATION.
RA Medigue C., Bocs S.;
RL Unpublished observations (DEC-2001).
CC -!- SIMILARITY: Belongs to the S17E family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
```

```

CC -----
DR EMBL; AP000066; ; NOT_ANNOTATED_CDS.
DR HAMAP; MF 00511; ; 1.
DR InterPro; IPR001210; Ribosomal_S17E.
DR Pfam; PF00833; Ribosomal_S17E; 1.
DR PROSITE; PS00712; RIBOSOMAL_S17E; 1.
DR Ribosomal protein; Complete proteome.
SQ SEQUENCE 67 AA; 8039 MW; D6245539F9835AD3 CRC64;

Query Match 7.5%; Score 5; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 NKKKV 59
DB 31 NKKKV 35

RESULT 14
CD14_LITXA
ID CD14_LITXA STANDARD; PRT; 12 AA.
AC P56246;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Caeridin 1.4.
OS Litoria xanthomera (Orange-thighed frog), and
OS Litoria chloris (Blue-thighed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae;
OC Pelodyadinae; Litoria.
OC NCBI_TaxID=79697, 86064;
CX NCBI_TaxID=79697, 86064;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC SPECIES=L.xanthomera.
RX MEDLINE=97374000; PubMed=9230483;
RA Steinborner S.T., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.,
RA Ramsay S.L.;
RT "New caerin antibacterial peptides from the skin glands of the
RT Australian tree frog Litoria xanthomera.";
RL J. Pept. Sci. 3:181-185(1997).
RN [2]
RP SEQUENCE.
RC SPECIES=L.chloris; TISSUE=Skin secretion;
RX MEDLINE=98175802; PubMed=9516047;
RA Steinborner S.T., Currie G.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "New antibiotic caerin 1 peptides from the skin secretion of the
RT Australian tree frog Litoria chloris. Comparison of the activities of
RT the caerin 1 peptides from the genus Litoria.";
RL J. Pept. Res. 51:121-126(1998).
CC -!- FUNCTION: Caeridins show neither neuropeptide activity nor
CC antibiotic activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC -!- MASS SPECTROMETRY: MW=1096; METHOD=FAB.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1097 MW; 28225503E3772728 CRC64;

Query Match 6.0%; Score 4; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGLG 7
DB 8 GGLG 11

RESULT 15
UKA2_HUMAN
ID UKA2_HUMAN STANDARD; PRT; 12 AA.
AC P31144;
DT 01-JUL-1993 (Rel. 26, Created)

```

```

DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of epidermal keratinocytes (spot 1617)
DE (Fragments).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.93, its MW is: 81.6 kDa.
DR Aarhus/Ghent-2DPAGE; 1617; IEF.
FT NON_TER 1 1
FT NON_CONS 7 8
FT UNSURE 8 8
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1351 MW; D6CD4A5E75F2C1F6 CRC64;

Query Match 6.0%; Score 4; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LGGR 9
DB 9 LGGR 12

Search completed: April 1, 2004, 10:05:46
Job time : 15 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 10:03:19 ; Search time 38 Seconds
(without alignments)
556.308 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 67

Sequence: 1 GFGGLGGRKCPSEIFSR.....CRLGYLRNKKVCVPRSKG 67

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 80005

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Listing first 45 summaries

Database : SPREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	10.4	58	11 Q99NG9	Q99ng9 mus musculu
2	6	9.0	49	4 Q86WB9	Q86wb9 homo sapien
3	6	9.0	64	16 Q8DW99	Q8dw99 streptococc
4	5	7.5	19	12 Q85671	Q85671 reovirus (t
5	5	7.5	20	12 Q85670	Q85670 reovirus (t
6	5	7.5	23	11 Q99LQ2	Q99lq2 mus musculu
7	5	7.5	25	4 Q14497	Q14497 equus caball
8	5	7.5	28	8 Q9T2U2	Q9t2u2 bos taurus
9	5	7.5	30	16 Q98AK1	Q98ak1 rhizobium l
10	5	7.5	30	16 Q835H6	Q835h6 enterococcu
11	5	7.5	32	16 Q9K175	Q9k175 neisseria m
12	5	7.5	34	6 Q9N1V1	Q9n1v1 equus caball
13	5	7.5	35	11 Q7TPO0	Q7tpo0 mus musculu
14	5	7.5	36	16 Q8E214	Q8e214 streptococc
15	5	7.5	37	8 Q78854	Q78854 phytophthor
16	5	7.5	37	16 Q7VKX2	Q7vqx2 haemophilus

17	5	7.5	38	5 P91965	P91965 penaeus van
18	5	7.5	40	2 Q48641	Q48641 lactococcus
19	5	7.5	40	5 Q86PP9	Q86pp9 culex pipie
20	5	7.5	41	4 Q9UDP7	Q9udp7 homo sapien
21	5	7.5	42	6 Q9GKT1	Q9gkt1 sus scrofa
22	5	7.5	42	10 Q23471	Q23471 arabidopsis
23	5	7.5	43	8 Q9TUN5	Q9tjns toxoplasma
24	5	7.5	43	12 Q65289	Q65289 human adeno
25	5	7.5	44	16 Q8F413	Q8f413 leptospira
26	5	7.5	44	16 Q7UGK1	Q7ugk1 rhodopirell
27	5	7.5	46	13 Q42309	Q42309 cyprinus ca
28	5	7.5	47	2 Q9AM66	Q9ame6 xanthomonas
29	5	7.5	47	7 Q30105	Q30105 homo sapien
30	5	7.5	47	10 Q98857	Q98857 glycine max
31	5	7.5	48	11 Q60413	Q60413 cricetus cr
32	5	7.5	49	4 Q8IZW1	Q8izw1 homo sapien
33	5	7.5	49	16 Q7UWE9	Q7uwe9 rhodopirell
34	5	7.5	51	2 Q31051	Q31051 streptococc
35	5	7.5	51	12 Q8JP04	Q8jpo4 feline coro
36	5	7.5	51	12 Q8JP05	Q8jpo5 feline coro
37	5	7.5	51	16 Q7UAP3	Q7uap3 shigella fl
38	5	7.5	53	9 Q854P4	Q854p4 mycobacteri
39	5	7.5	53	16 Q8ZIQ8	Q8ziqu8 yersinia pe
40	5	7.5	53	16 Q8CLS8	Q8cls8 yersinia pe
41	5	7.5	54	16 Q8XCC4	Q8xcc4 escherichia
42	5	7.5	56	4 Q8IVM4	Q8ivm4 homo sapien
43	5	7.5	56	11 Q9QWE6	Q9qwe6 rattus sp.
44	5	7.5	56	16 Q8XQWS	Q8xqws ralstonia s
45	5	7.5	58	4 Q9BYX8	Q9byx8 homo sapien

ALIGNMENTS

RESULT 1

ID	Q99NG9	PRELIMINARY;	PRT;	58 AA.
AC	Q99NG9;			
DT	01-JUN-2001	(TrEMBLrel. 17, Created)		
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	High-glycine/tyrosine protein type I ES.			
GN	AY026312.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C3H/HeN; TISSUE=Skin;			
RA	Fujiwara K., Poirier C., Hattori A., Noro C., Yoshiki A., Kusakabe M.;			
RT	"cDNA cloning of a new member of mouse high-glycine/tyrosine protein."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Lung;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The RANTOM Consortium,			
RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."			
RL	Nature 420:563-573(2002).			
DR	EMBL; AY026312; AAK07673.1; -.			
DR	EMBL; AK086935; BAG39767.1; -.			
DR	MGD; MGI:2181750; AY026312.			
DR	InterPro; IPR001064; Crystallin.			
DR	PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.			
SQ	SEQUENCE 58 AA; 6031 MW; 6FD440F5F9918C09 CRC64;			

Query Match 10.4%; Score 7; DB 11; Length 58;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LCIKIC 42
Db 45 LCIKIC 50

RESULT 4
Q85671 PRELIMINARY; PRT; 19 AA.
AC Q85671; TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE Activation-induced cytidine deaminase (Fragment).
GN AID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nishimura M., Yabe H., Shimada E., Mitsunaga S.,
RA Akaza T., Juji T.;
RT "A novel missense mutation (R24Q) in the AID (activation-induced
RT cytidine deaminase) gene in a Japanese patient with hyper-IGM
RT syndrome.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB092577; BAC66114.1; -.
FT NON_TER 1
FT NON_TER 49
SQ SEQUENCE 49 AA; 6068 MW; 8DC55D224EA186E7 CRC64;

Query Match 9.0%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GYLK 56
Db 44 GYLK 49

RESULT 3
Q8DW99 PRELIMINARY; PRT; 64 AA.
AC Q8DW99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
GN SMU168.
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AB014868; AAN57944.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH 3.
DR Pfam; PF01381; HTH 3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Complete proteome.
SQ SEQUENCE 64 AA; 7105 MW; 219453D6A513C7AE CRC64;

Query Match 9.0%; Score 6; DB 16; Length 64;

Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LCIKIC 42
Db 45 LCIKIC 50

RESULT 4
Q85671 PRELIMINARY; PRT; 19 AA.
AC Q85671; TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE Activation-induced cytidine deaminase (Fragment).
GN AID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nishimura M., Yabe H., Shimada E., Mitsunaga S.,
RA Akaza T., Juji T.;
RT "A novel missense mutation (R24Q) in the AID (activation-induced
RT cytidine deaminase) gene in a Japanese patient with hyper-IGM
RT syndrome.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB092577; BAC66114.1; -.
FT NON_TER 1
FT NON_TER 49
SQ SEQUENCE 49 AA; 6068 MW; 8DC55D224EA186E7 CRC64;

Query Match 9.0%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GYLK 56
Db 44 GYLK 49

RESULT 3
Q8DW99 PRELIMINARY; PRT; 64 AA.
AC Q8DW99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
GN SMU168.
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AB014868; AAN57944.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH 3.
DR Pfam; PF01381; HTH 3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Complete proteome.
SQ SEQUENCE 64 AA; 7105 MW; 219453D6A513C7AE CRC64;

Query Match 9.0%; Score 6; DB 16; Length 64;

Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LCIKIC 42
Db 45 LCIKIC 50

RESULT 4
Q85671 PRELIMINARY; PRT; 19 AA.
AC Q85671; TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE Activation-induced cytidine deaminase (Fragment).
GN AID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nishimura M., Yabe H., Shimada E., Mitsunaga S.,
RA Akaza T., Juji T.;
RT "A novel missense mutation (R24Q) in the AID (activation-induced
RT cytidine deaminase) gene in a Japanese patient with hyper-IGM
RT syndrome.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB092577; BAC66114.1; -.
FT NON_TER 1
FT NON_TER 49
SQ SEQUENCE 49 AA; 6068 MW; 8DC55D224EA186E7 CRC64;

Query Match 9.0%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GYLK 56
Db 44 GYLK 49

RESULT 3
Q8DW99 PRELIMINARY; PRT; 64 AA.
AC Q8DW99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
GN SMU168.
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AB014868; AAN57944.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH 3.
DR Pfam; PF01381; HTH 3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Complete proteome.
SQ SEQUENCE 64 AA; 7105 MW; 219453D6A513C7AE CRC64;

Query Match 9.0%; Score 6; DB 16; Length 64;

Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 LCIKIC 42
Db 45 LCIKIC 50

RESULT 4
Q85671 PRELIMINARY; PRT; 19 AA.
AC Q85671; TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE Activation-induced cytidine deaminase (Fragment).
GN AID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe Y., Nishimura M., Yabe H., Shimada E., Mitsunaga S.,
RA Akaza T., Juji T.;
RT "A novel missense mutation (R24Q) in the AID (activation-induced
RT cytidine deaminase) gene in a Japanese patient with hyper-IGM
RT syndrome.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB092577; BAC66114.1; -.
FT NON_TER 1
FT NON_TER 49
SQ SEQUENCE 49 AA; 6068 MW; 8DC55D224EA186E7 CRC64;

Query Match 9.0%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 GYLK 56
Db 44 GYLK 49

RESULT 3
Q8DW99 PRELIMINARY; PRT; 64 AA.
AC Q8DW99;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator.
GN SMU168.
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AB014868; AAN57944.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH 3.
DR Pfam; PF01381; HTH 3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Complete proteome.
SQ SEQUENCE 64 AA; 7105 MW; 219453D6A513C7AE CRC64;

Query Match 9.0%; Score 6; DB 16; Length 64;

```

DR InterPro: IPR004317; Sigma.1.2.
FR Pfam: PF03084; Sigma.1.2; I.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2165 MW; A834D093E79852F7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Length 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GFGGL 6
Db 12 GFGGL 16

RESULT 6
Q99LQ2 PRELIMINARY; PRT; 23 AA.
AC Q99LQ2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; BC002276; AAH02276.1; -.
KW Hypothetical protein.
SQ SEQUENCE 23 AA; 2223 MW; 2A4C872213B58B01 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 4.3e+02; Length 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGLGG 8
Db 19 GGLGG 23

RESULT 7
Q14497 PRELIMINARY; PRT; 25 AA.
AC Q14497;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CCAAT binding factor subunit C (Fragment).
GN HCBF-C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=brain;
RC Dittus=brain;
RA Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I.,
RA Kavsan V.M.;
RT "Characterization of different mRNA types expressed in human brain.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR ENBL; Z70024; CAA93846.1; -.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2649 MW; 12DE150C48C31875 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 4.6e+02; Length 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGG 5
Db 19 GFGGG 5

RESULT 8
Q9T2U2 PRELIMINARY; PRT; 28 AA.
AC Q9T2U2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE NADH:ubiquinone oxidoreductase (Complex I) iron-sulfur protein
DE fraction 20 kDa polypeptide N-terminus (Fragment).
OS Bos taurus (Bovine).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=92138662; PubMed=1778979;
RA Masui R., Wakabayashi S., Matsubara H., Hatemi Y.;
RT "The amino acid sequence of the 9 kDa polypeptide and partial amino
RT acid sequence of the 20 kDa polypeptide of mitochondrial
RT NADH:ubiquinone oxidoreductase.";
RL J. Biochem. 110:575-582(1991).
FT NON_TER 1
FT NON_TER 28
SQ SEQUENCE 28 AA; 2852 MW; 8C86A28689F027C8 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 5.1e+02; Length 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 VVPXP 36
Db 14 VVPXP 18

RESULT 9
Q98AK1 PRELIMINARY; PRT; 30 AA.
AC Q98AK1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein mar5969.
GN MSR5969.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Igesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR ENBL; AP003008; BAB52328.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3375 MW; 33769641EBB4CBF3 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 5.4e+02; Length 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 RLGYL 53

```

```

Db      12 RLGYL 16
|||||
RESULT 10
Q835H6 PRELIMINARY; PRT; 30 AA.
ID Q835H6
AC Q835H6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
CN EF1401.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12633927;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA Dougherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson M.,
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,
RA "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis."
RL Science 299:2071-2074(2003).
DR EMBL; AE016951; AAC01192.1; -.
DR TIGR; EF1401; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 30 AA; 3449 MW; 979FCACE415AF8BD CRC64;

Query Match 7.5%; Score 5; DB 16; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 VPRSK 65
|||||
Db 10 VPRSK 14

RESULT 11
Q9K175 PRELIMINARY; PRT; 32 AA.
ID Q9K175
AC Q9K175
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein NMB0297.
CN NMB0297.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson M.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masiagnani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002386; AAF40748.1; -.
DR PIR; H81215; H81215.

```

```

DR TIGR; NMB0297; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 32 AA; 3563 MW; 988618AE929A5D8D CRC64;

Query Match 7.5%; Score 5; DB 16; Length 32;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GGRGK 11
|||||
Db 20 GGRGK 24

RESULT 12
Q9NIV1 PRELIMINARY; PRT; 34 AA.
ID Q9NIV1
AC Q9NIV1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Prion protein (Fragment).
CN PRNP.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20082971; PubMed=10613847;
RA Caetano A.R., Shiue Y.L., Lyons L.A., O'Brien S.J., Laughlin T.F.,
RA Bowling A.T., Murray J.D.;
RT "A comparative gene map of the horse (Equus caballus).";
RL Genome Res. 9:1239-1249(1999).
DR EMBL; AF134232; AAF62346.1; -.
DR InterPro; IPR000817; Prion.
DR PROSITE; PS00291; PRION_1; 1.
FT NON_TER 1 34
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 3212 MW; A75D6E78E4DC0075 CRC64;

Query Match 7.5%; Score 5; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGLGG 8
|||||
Db 24 GGLGG 28

RESULT 13
Q7TPH0 PRELIMINARY; PRT; 35 AA.
ID Q7TPH0
AC Q7TPH0
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FGF2 (Fragment).
CN FGF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RA Foletti A., Vuadenes P., Beermann F.;
RT "Intracellular localization of mouse FGF2."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY324448; AAP92384.1; -.
FT NON_TER 35 35
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3451 MW; 08F3AF21ABFC7250 CRC64;

Query Match 7.5%; Score 5; DB 11; Length 35;

```

Best Local Similarity 100.0%; Pred. No. 6.2e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGGRG 10
|
|
|
|
|
Db 9 LGGRG 13

RESULT 14
Q8E214 PRELIMINARY; PRT; 36 AA.
AC Q8E214;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SAG0186.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan A.M., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AF014200; AM99093.1; -;
DR TIGR; SAG0186; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 36 AA; 4411 MW; 9072B5F8C03CE2C0 CRC64;

Query Match 7.5%; Score 5; DB 16; Length 36;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 NKKKV 59
|
|
|
|
|
Db 13 NKKKV 17

RESULT 15
O78854 PRELIMINARY; PRT; 37 AA.
AC O78854;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ribosomal protein (Fragment).
GN RPL2.
OS Phytophthora cinnamomi.
OG Mitochondrion.
OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID=4785;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2400;
RX MEDLINE=98336900; PubMed=9673029;
RA Dobrowolski M.P., Tommerup I.C., O'Brien P.A.;
RA "Microsatellites in the mitochondrial genome of phytophthora cinnamomi
RT failed to provide highly polymorphic markers for population
RT genetics."

RL FEMS Microbiol. Lett. 163:243-248 (1998).
DR EMBL; AF051358; AAC39491.1; -;
DR GO; GO:0005822; C:intracellular; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR002171; Ribosomal_L2.
DR InterPro; IPR008991; Transl_SH3_like.
DR Pfam; PF03947; Ribosomal_L2_C; 1.
DR KW Mitochondrion.
FT NON TER
SQ SEQUENCE 37 AA; 3862 MW; 487EA067E18B1EE0 CRC64;
Query Match 7.5%; Score 5; DB 8; Length 37;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 RNKKK 58
|
|
|
|
|
Db 27 RNKKK 31

Search completed: April 1, 2004, 10:07:11
Job time : 41 secs